



3R-Red de Repositorios Universitarios de Recursos Digitales

Etapas 2: Modelo Conceptual Segundo Informe Técnico

Responsables:

Clara López Guzmán – DGSCA (Coordinadora)
Alberto Arriaga Arredondo - DGSCA
Alberto Castro Thompson – DGB
Isabel Galina Russell – DGSCA
Fernando Gamboa Rodríguez – CCADET
Joaquín Giménez Heau - Instituto de Biología
Pablo Miranda Quevedo – DGSCA

Enero, 2007

1. INTRODUCCIÓN	3
2. LOS REPOSITARIOS UNIVERSITARIOS	4
2.1. Definición	4
2.2. Operación	4
2.2.1. Usuarios	5
2.2.2. Contenidos	5
2.2.3. Políticas	6
2.2.4. Costos	7
2.3. Modelos para la interconexión de RU	8
2.3.1. Repositorio centralizado	9
2.3.2. Repositorio distribuido	9
2.3.3. Repositorio híbrido	10
3. EVALUACIÓN DE HERRAMIENTAS	11
3.1. DSpace	11
3.2. FEDORA	13
3.3. Propuesta de solución	13
4. MODELO CONCEPTUAL DE 3R	15
4.1. Operación general	15
4.2. Contenidos	17
4.2.1. Tipos de objetos digitales	18
4.2.2. Proceso de validación	19
4.2.3. Integración de colecciones	20
4.3. Políticas	22
4.3.1. Políticas sobre la gestión de contenidos y colecciones	22
4.3.2. Políticas sobre procedimientos de depósito	24
4.3.3. Políticas de uso	24
4.4. Arquitectura de un repositorio local	25
5. CONCLUSIONES Y TRABAJO FUTURO	27
6. REFERENCIAS	28
7. BIBLIOGRAFÍA	29
8. ANEXOS	31

1. Introducción

El proyecto 3R-Red de Repositorios Universitarios de Recursos Digitales¹, forma parte del Macroproyecto de Tecnologías para la Universidad de la Información y la Computación. Es un proyecto que consta de cuatro etapas: Investigación, Modelo Conceptual, Desarrollo e Implementación.

El objetivo del proyecto está encaminado a la creación del prototipo de una red de repositorios de la UNAM, que permita mayor uso y visibilidad de la producción intelectual de los miembros de la comunidad.

Este informe técnico presenta los resultados de la segunda etapa del proyecto, Modelo Conceptual, en la que se realizó un trabajo multidisciplinario de discusión para obtener un primer modelo conceptual, que sirva de base para la etapa de desarrollo del prototipo de la red. Este modelo proporciona información, de manera gráfica, acerca de qué es lo que hace el prototipo 3R y los mecanismos que implementará para llevar a cabo sus tareas y cumplir sus objetivos. Se obtuvo a partir de la investigación que se realizó en la etapa de investigación previa, sobre los conceptos básicos y herramientas en torno a los repositorios, así que esos conceptos se dan por conocidos y para recordarlos se debe recurrir al documento técnico de la primera etapa de 3R.

En el informe de la primera etapa, Investigación, se abundó en el tema de los repositorios, en conocer sus características, las tecnologías involucradas y las herramientas disponibles para su conformación. Asimismo se llegó a establecer que el tipo de repositorios que se manejarán en la UNAM, dentro del proyecto 3R, se denominarán repositorios universitarios (RU). Así que, a partir del trabajo realizado en la primera etapa y acorde con las características específicas de la UNAM, se ha determinado que un RU es: *un depósito de contenidos digitales informativos, educativos y de investigación, con un conjunto de servicios en línea para su difusión, uso y visibilidad, producido y administrado por miembros de la comunidad académica de la UNAM, de forma institucional o grupal.*

Partiendo de esta definición, en el estudio realizado en esta segunda etapa del proyecto, se ha procedido a determinar la tipología, la infraestructura y las políticas generales que se seguirán para dar marcha a la conformación de los RU y de la red 3R. También se hace un planteamiento inicial de la forma en la que deben operar los RU, tanto de forma local como dentro de 3R.

En el siguiente apartado de este documento se presentan los conceptos generales entorno a los repositorios universitarios; después, en el apartado tres, se hace un análisis detallado de las herramientas de software que se están considerando adecuadas para el modelo. En el apartado cuatro, se describe con detalle el modelo conceptual de 3R; en el apartado cinco se presentan algunas conclusiones del trabajo realizado y, finalmente, se integran las referencias, la bibliografía y los anexos correspondientes.

¹ <http://www.3r.unam.mx/>

2. Los repositorios universitarios

El modelo de repositorio que se plantea en 3R surge a raíz del estudio de los tipos y modelos de repositorios que actualmente se están gestando en todo el mundo. A partir de éstos se ha integrado una propuesta que satisface los requisitos de los repositorios de la UNAM. A continuación se detallan algunos rubros importantes para el entendimiento de este tipo de repositorios, que dentro de 3R, se reconocen como repositorios universitarios.

2.1. Definición

En el primer entregable de este proyecto se documentó que existen diferentes tipos de repositorio que responden a diferentes necesidades de manejo de información. En ese primer informe se mencionaba que a nivel mundial los repositorios se pueden clasificar en:

- Repositorios de eprints y temáticos. Contienen documentos de un mismo tema, usualmente artículos científicos.
- Repositorios de materiales académicos. No contiene sólo documentos científicos arbitrados, sino que alberga todo tipo de materiales que apoyen la enseñanza y el aprendizaje, que pueden o no corresponder al mismo tema.
- Repositorios institucionales. Incluyen material académico diverso, tienden a ser organizados por una institución más que por áreas temáticas.

Dada la amplia producción intelectual de la UNAM y la diversidad de recursos que esto genera, así como la complejidad administrativa que una universidad de grandes dimensiones implica, no es fácil encontrar una relación inmediata entre los tipos de repositorios que actualmente existen y los diversos materiales que aquí se producen. Por ello, se ha desarrollado el término *repositorio universitario (RU)* que describe, como ya de vio arriba, el tipo de contenido que puede adaptarse a los objetivos de 3R, entendiéndose a un RU como: un depósito de contenidos digitales informativos, educativos y de investigación, con un conjunto de servicios en línea para su difusión, uso y visibilidad, producido y administrado por la comunidad académica de la UNAM, de forma institucional o grupal.

2.2. Operación

Los RU son, en su concepción más simple, sistemas de información que permiten el manejo y gestión de datos con el fin de proveer información a usuarios que tengan necesidad de un acceso sencillo y seguro a los recursos que faciliten la adquisición de conocimientos.

Como todo sistema de información, los repositorios universitarios deben contar con elementos que permitan su funcionamiento y cumplir su objetivo: facilitar los contenidos a quien así lo requiera y facilitar la administración de los mismos a los productores. Se plantea que un RU debe considerar para su operación:

- a) Usuarios
- b) Contenidos
- c) Políticas
- d) Costos
- e) Herramientas de software

En los siguientes incisos se abordarán los rubros de usuarios, contenidos, políticas y costos. El tema de herramientas se abordará en el apartado tres.

2.2.1. Usuarios

Por la forma en que operan los sistemas de información, se tienen dos grandes grupos de usuarios:

- **Usuarios internos:** son usuarios que tienen una estrecha relación con el sistema en cuanto a su alimentación, actualización y buen funcionamiento en general; ya sea porque se trata de personal que se dedica a la gestión o al mantenimiento del sistema o porque se trate de los proveedores de la información que será manipulada por el mismo.
- **Usuarios externos:** entidades cuya relación con el sistema se limita a la consulta de la información contenida en el mismo.

El caso de los repositorios universitarios no es la excepción, por lo que adoptan estos mismos tipos de usuarios, de la siguiente forma:

- **Usuarios internos:** administradores, responsables tecnológicos, responsables de contenidos, proveedores de objetos digitales.
- **Usuarios externos:** personas que consultan los contenidos de cada uno de los repositorios por separado, o desde una interfaz común acceden a todos los repositorios.

Además de esta distinción de tipos de usuarios se tienen dos niveles:

- **Usuarios a nivel local:** usuarios de un solo repositorio universitario.
- **Usuarios a nivel general:** usuarios de un conjunto de repositorios (seguramente en red) o de un nodo central que conecte a un grupo de repositorios.

Entiéndase que dentro del rubro de *usuarios locales y generales* se consideran tanto administradores como aquellos que hacen consultas al repositorio de manera interna y/o externa.

2.2.2. Contenidos

Los tipos de documentación que se utilizan en las actividades universitarias de educación, investigación y extensión de la cultura son variados, además de que dicha documentación se presenta en una amplia variedad de formatos digitales.

No es una labor trivial identificar todos los posibles tipos de materiales, recursos o documentos digitales utilizados para representar y conservar la información. Sin embargo, debe partirse de identificar un conjunto de estas representaciones o tipos de contenidos, a fin de normalizarlos dentro del repositorio. Esto facilita el establecer planes de acción para el caso de que se tenga contemplada la preservación de la información o el simple mantenimiento del repositorio en un largo periodo de tiempo.

A continuación se muestra una primera lista con los tipos de contenido que se se pueden llegar a utilizar en el contexto universitario (estos contenidos se extrajeron en base a los

requisitos de evaluación de productividad académica de la UNAM y de los recursos digitales que se encontraron mencionados con mas frecuencia en los estudios de RIs internacionales y nacionales realizados para la Etapa 1 de este proyecto):

- Artículos con o sin arbitraje en revistas nacionales e internacionales
- Libros
- Traducciones de libros
- Bases de datos
- Monografías
- Memorias arbitradas en congresos nacionales o internacionales
- Proyectos
- Reportes técnicos
- Manuales
- Programas de cómputo
- Programas académicos
- Apuntes de cursos
- Recursos educativos
- Multimedios (audio, videos, modelos VR y publicaciones interactivas)
- Conferencias
- Ponencias
- Cursos
- Pósters
- Informes
- Documentos administrativos de proyectos

Cabe resaltar la importancia de que el soporte en el que sean grabados estos contenidos sea apto para su almacenamiento en un servidor en línea, y muy deseablemente que el formato sea compatible con la tecnología web, aunque no es una restricción.

2.2.3. Políticas

Las políticas son el conjunto de reglas normativas que define y formaliza la operación y la organización de los repositorios en sus distintos niveles de funcionamiento y servicios; establecen las tareas y responsabilidades de las partes involucradas en su gestión administración y en el uso de los contenidos.

De forma genérica, las políticas se centran en aspectos relativos a: tipos de contenido, organización de colecciones, ingreso de materiales, validación de metadatos, usuarios y derechos legales, entre otros.

La adopción de políticas y reglas trae consigo beneficios que impactan en la facilidad de utilización del sistema, en la organización de contenidos y en la interoperabilidad entre repositorios. Asimismo, facilitan el trabajo a usuarios y a administradores ya que proveen reglas claras para realizar operaciones y procedimientos, tanto en el repositorio, como en sus servicios y aplicaciones.

Finalmente es importante remarcar que las políticas deben, a su vez, apegarse a las normativas de la institución, así como a la idiosincrasia de la comunidad que genera y utiliza los repositorios, siempre con la intención de seguir un orden en los procesos,

mantener un control sobre los contenidos y los usuarios y facilitar el acceso y operación de los sistemas.

2.2.4. Costos

Si bien los repositorios universitarios ofrecen sus contenidos gratuitamente, estructurarlos y ponerlos en funcionamiento, desde el punto de vista técnico y organizacional, conlleva gastos que deben ser considerados en las distintas etapas del desarrollo como la etapa de pruebas y gestación. Un repositorio, aun cuando ya esta establecido en el marco de una arquitectura estable, requiere de personal dedicado tiempo completo para una continua, detallada y cuidadosa gestión.

Para estimar los costos de la organización de los repositorios, se deben tener en cuenta distintos elementos que intervienen en sus diferentes etapas de desarrollo y modalidades de estructuración, como:

- a) Infraestructura mínima por repositorio, costo de los equipos básicos (servidores, recolectores y proveedores de datos).
- b) Costo del soporte técnico e instalación, que aunque devenga de plataformas de código abierto, implican la adaptación local del software que puede requerir varios meses para el desarrollo del sistema y la interfaz.
- c) Interconexión con servicios externos y otras redes.
- d) Administración, organización y desempeño del repositorio. Requiere de un grado alto en la especialización e incorporación de personal con habilidades muy específicas como programadores, diseñadores web y conocimientos de distintos esquemas de metadatos para su validación y control de calidad, y personal dedicado a la gestión y organización de la información autodepositada, y a la gestión institucional, con la comunidad de investigadores y los mandos jerárquicos de la institución y los departamentos. Los costos del staff deben ser considerados a largo plazo con un presupuesto destinado a mantener los equipos que hacen funcionar los repositorios, sean centralizados o distribuidos.
- e) Validez y calidad de contenidos; el costo de la gestión de los metadatos en función del número de depósitos que recibe un repositorio o cuantos podría recibir potencialmente y su consiguiente control de calidad y validación de los metadatos incorporados; en ocasiones los RU pueden incorporar material preexistente de otras colecciones, lo cual representa procesos de migración de metadatos originales de mayor complejidad y con un alto nivel de especialización.
- f) Capacitación, consiste en tener el personal adecuado que pueda brindar conocimientos sobre el funcionamiento del repositorio principalmente a dos niveles, en el primero a los involucrados en el funcionamiento básico de los repositorios locales: responsables técnicos, editores de metadatos y gestores; en el segundo tiene que ver con la sensibilización de los núcleos iniciales de las comunidades académicas, a las cuales se deben transmitir nociones sobre los

beneficios de los repositorios, así como las formas en las cuales se utilizan los recursos y herramientas implementados.

- g) Preservación y actualización tecnológica. Actualmente la preservación digital es un terreno poco conocido, sus modelos aún están en proceso de definición y experimentación lo cual hace que no haya elementos suficientes para poder definir sus costos. Por esta razón este rubro no debe ser contemplado en etapas posteriores de la planeación de los repositorios universitarios y debe considerarse en los presupuestos iniciales. Esta actividad tiene como tareas básicas y concretas: el almacenamiento, migración y actualización permanente de los estándares operativos que permitan la consulta e interoperabilidad de los metadatos y objetos depositados a lo largo del tiempo y requiere de equipos y software dedicados así como de personal especializado. Para propósitos de planeación a largo plazo, algunas fuentes sugieren destinar el 10 % del presupuesto (Jones et. al. 2006) general destinado a los repositorios universitarios.

Un elemento importante para establecer un repositorio que funcione a largo plazo, es el costo del tiempo de gestión y organización, es decir, el tiempo y esfuerzo necesario para acomodar los elementos que consoliden los progresos hechos en los distintos ámbitos de la organización del mismo, conformar un equipo humano apropiado, productivo y motivado, que establezca y ajuste las políticas de funcionamiento y vinculación con las comunidades académicas y las autoridades universitarias y que tenga suficiencia en el manejo técnico de las herramientas y flujos de trabajo respecto a los metadatos, y su preservación.

Asimismo, los RU deben establecer un monto estimado acorde a la producción promedio de su planta de académicos, y a la progresión de los depósitos que se podrán efectuar en periodos de tiempo regulares (Jones et. al. 2006), de esta manera ajustar el presupuesto a las necesidades de almacenamiento, cuidado de metadatos y mantenimiento del staff necesario para su funcionamiento.

En esta etapa de la investigación o se obtuvo un modelo para estimar estos costos, dado que la experiencia real se tendrá en la etapa siguiente y será en ésta en la que se haga una propuesta al respecto.

2.3. Modelos para la interconexión de RU

Hasta este punto se ha dado una idea global de lo qué es un RU y sus componentes principales. Sin embargo el proyecto 3R tiene como objetivo la creación de una red; para ello, es necesaria la interconexión de diversos RU.

La organización de los componentes y formas de funcionamiento de los RU puede plantearse de acuerdo a tres modelos de arquitecturas: centralizada, distribuida e híbrida. Cada una de ellas requiere formas de organización distintas respecto a los costos de la infraestructura técnica, la especialización del personal y la adopción de políticas de administración y preservación. A continuación se explican los modelos a fin de justificar el modelo que se presenta en el apartado cuatro.

2.3.1. Repositorio centralizado

El repositorio centralizado concentra en un solo servidor los contenidos depositados y los servicios de acceso y preservación, tiene como ventajas que puede establecer un servicio de administración general que controle la totalidad de la estructura de las colecciones y los mecanismos de control de calidad sobre los metadatos, así como la posibilidad de supervisar y reponer rápidamente cualquier parte del sistema cuando se presente alguna falla.

La infraestructura mínima requiere poco equipo pero de gran capacidad y alto rendimiento y esto hace más simple el mantenimiento y actualización tecnológica; requiere de poco personal pero con un grado mayor de especialización en la administración del sistema y en la organización del repositorio, un grupo de trabajo y la conformación de un comité editorial, el volumen de trabajo y de procesos implicados (validez y calidad de contenidos) hacen que estas tareas sean mas complejas que en otros modelos. Respecto al resguardo y preservación la estructura es robusta y con un alto nivel de control, los costos generales son altos ya que recaen en la organización que implementa el repositorio y tienen que hacerse de una sola vez (ver anexo A).

2.3.2. Repositorio distribuido

Esta constituido por los repositorios instalados en cada una de las organizaciones o dependencias que los administran, cuentan con su propia infraestructura técnica y se rigen por políticas administrativas y de preservación particulares. A nivel local cuentan con sus propios mecanismos de consulta y depósito, al mismo tiempo se integran una red en la cual por medio de un portal de búsqueda y recuperación ofrecen sus metadatos y recursos.

La organización de este tipo de repositorio requiere de una mayor complejidad en los aspectos técnicos, de gestión y capacitación de los recursos humanos, ya que debe contar con distintos responsables en lo que respecta al área tecnológica y el área de contenidos y estructuración de sus colecciones, bien llevado este tipo de repositorio pueden dar un mejor desempeño y en caso de alguna falla no se afecta la totalidad del sistema, conlleva también ciertas dificultades respecto a la estandarización en el control de los metadatos y la autenticación de usuarios y los procesos de provisión de recursos para su cosecha.

Por otra parte el proceso de preservación y resguardo a largo plazo en este modelo plantea un mayor costo y problemática, ya que difícilmente una sola dependencia podrá sostener una política de resguardo vigente, actualizada y consistente.

Respecto a la infraestructura tecnológica requerida las inversiones iniciales pueden ser de bajo costo, pero se incrementan a largo plazo en lo que respecta el mantenimiento y operación de los distintos equipos de trabajo (ver anexo B).

2.3.3. Repositorio híbrido

El repositorio híbrido es una conjunción entre los dos modelos antes mencionados, es decir con una organización distribuida y un mecanismo de cosecha de metadatos y preservación centralizada. La forma de acceso puede ser local o a través de un portal común que extrae sus datos de un servidor caché que los cosecha de los repositorios asociados

Por una parte este modelo permite una estandarización general en los métodos de depósito, control y validación de metadatos, pero con las políticas administrativas y estructura de colecciones particulares de cada dependencia, así mismo la preservación se hace más operativa y segura ya que el mismo sistema deposita respaldos de la información en un contenedor centralizado.

Este modelo requiere de la instrumentación de políticas administrativas a dos niveles, por un parte las que se determinan a nivel local en las dependencias en sus distintos repositorios y por otra de un marco más amplio e institucional respecto al uso consulta, depósito y derechos de los recursos propiedad de la universidad.

Respecto al soporte tecnológico que requiere la red de este tipo de repositorios, los costos se reducen debido a que se puede constituir un solo equipo de trabajo y soporte que responda a las contingencias que se pueden presentar en los distintos repositorios asociados, por otra parte se deben constituir equipos independientes para administrar localmente los recursos y hacer un consejo general que regule la administración de todo el sistema (ver anexo C).

3. Evaluación de herramientas

En la etapa de investigación de proyecto se llevó a cabo una análisis de las herramientas de software más comunes para la creación y mantenimiento de repositorios, esto con la intención de tener fundamentos para la selección de soluciones de software para la creación, administración y mantenimiento de los RU. Dado que la Universidad produce una gran cantidad de material, en diversos formatos y de diferentes tipos, y que cada Dependencia que conforma a la UNAM tiene necesidades muy particulares en el manejo de su producción intelectual, la elección de una herramienta no resulta trivial ya que hay diversidad de opciones y cada una tiene características, ventajas y desventajas propias.

Para elegir la herramienta pertinente, se tomaron los siguientes criterios:

- 1) Que estuviera desarrollada con software libre y código abierto, ya que esto permite la modificación de su código fuente y facilita la creación de aplicaciones que se adaptan exactamente a necesidades particulares de manejo y diversidad de contenidos sin tener que iniciar de cero un nuevo desarrollo de software.
- 2) Que la herramienta ya fuera probada o utilizada por grupos de trabajo similares al proyecto 3R.
- 3) Que la herramienta utilizada tecnología de punta y tuviera posibilidad de extender hacia la web semántica. Esto permitirá que los RU tengan una plataforma más estable respecto a la rapidez del cambio tecnológico.

Así que, después de analizar las herramientas con base en estos criterios se eligieron dos posibles soluciones: DSpace² y FEDORA³.

DSpace se eligió debido a que es un software muy utilizado alrededor de todo el mundo ya que es muy fácil de poner en marcha y tiene interfaces y flujos de trabajo modificables sin dedicar demasiado esfuerzo para ello.

Por otro lado, FEDORA ha tenido un fuerte impacto entre implementadores de repositorios debido a que posee una robustez y versatilidad en el manejo de colecciones digitales que poco software posee. El núcleo de software que es FEDORA permite que se desarrollen aplicaciones que satisfacen necesidades muy particulares.

Estas herramientas se explican a continuación con más detalle.

3.1. *DSpace*

Es una aplicación para gestión de acervos digitales de carácter general que acepta cualquier formato digital en el que se puedan encontrar los contenidos digitales. Está escrito en el lenguaje de programación java y utiliza algunas herramientas extras al código principal de DSpace, para ampliar su funcionalidad y hacer sencilla su implementación y gestión. Entre las herramientas que utiliza DSpace para su

² <http://dspace.org/implement/index.html>

³ <http://www.fedora.info/documents/WhitePaper/FedoraWhitePaper.pdf>

funcionamiento están PostgreSQL (o algún otro sistema manejador de bases de datos relacionales), Apache y Tomcat (o algún otro servidor web y servidor de aplicaciones java).

Este software se distribuye bajo la licencia BSD (*Berkley Software Distribution*) la cual permite desarrollo de aplicaciones comerciales derivadas de la modificación del software que se distribuya bajo dicha licencia. Un detalle que hay que mencionar es que no todos los componentes del software pueden estar amparadas por la licencia BSD por lo que se debe tener cuidado para sustituir la parte o partes de las cuáles no se tiene permiso para modificar y distribuir de manera comercial.

DSpace puede ejecutarse en cualquier plataforma que soporte JVM (*Java Virtual Machine*) incluyendo todas las variaciones de Unix y Linux, Mac OSX y Microsoft Windows. El sitio del proyecto DSpace en SourceForge.net proporciona todo el código fuente así como la documentación para poder instalar esta herramienta.

DSpace tiene como principales características funcionales (Jones et. al. 2006):

Depósito de contenido digital y sus metadatos (Dublin Core⁴) asociados en el repositorio por parte de usuarios (investigadores y académicos universitarios) siguiendo un flujo de trabajo definido de manera particular para cada colección que el repositorio maneja de manera interna; flujo que permite tener un control en la manera en cómo se revisan y aprueban los contenidos que estarán almacenados en el repositorio.

Gestión y manejo de los contenidos digitales del repositorio durante largos periodos de tiempo.

Búsqueda/navegación y recuperación de contenidos digitales por medio de una interfaz Web sencilla, esto incluye algún tipo de soporte para búsquedas avanzadas en metadatos estructurados o a texto completo.

Funciones de control de acceso que permiten restringir las personas que pueden alimentar al repositorio con nuevos objetos digitales; dichas funciones de permiten controlar el acceso a colecciones de objetos e incluso restringir el acceso a objetos individuales.

Manejo de licencias individuales para cada ítem almacenado en el repositorio que garanticen a los administradores del repositorio permisos explícitos para administrar, preservar y diseminar los contenidos.

Creciente comunidad de desarrolladores de software *open source* alrededor de todo el mundo que permite una mayor facilidad para encontrar soluciones a posibles problemas que se pueden presentar al implementar DSpace como software de repositorios.

⁴ <http://es.dublincore.org/>

3.2. FEDORA

La producción intelectual en la Universidad es diversa y se distribuye en muchos tipos de contenidos (tesis, work papers, reportes de investigación, artículos de divulgación, libros, monografías, imágenes, videos, archivos de sonido, etc.) y formatos (MS Word, MS Excel, MS PowerPoint, PDF, JPG, BMP, GIF, TIFF, PNG, MPEG, AVI, WMV, WMA, OGG, MP3, etc.). El software para repositorios FEDORA (Flexible Extensible Digital Object and Repository Architecture) fue desarrollado teniendo en cuenta que el contenido digital, resultado de la labor de investigación, no sólo son documentos de texto sino que existe una gran variedad de formatos que docentes e investigadores utilizan para difundir su trabajo. De manera general se puede dividir en dos el tipo de material que los administradores de repositorios tienen que gestionar:

Objetos convencionales: libros y otros objetos de tipo texto, datos geoespaciales, imágenes, mapas.

Objetos complejos y dinámicos: video, conjuntos de datos numéricos y sus códigos asociados, audio.

Conforme aumenta la diversidad de tipos de medios que conforman los acervos digitales de instituciones que se dedican a la investigación, los responsables de resguardar dichas colecciones se enfrentan a tareas cada vez más complejas. FEDORA intenta dar respuesta a cuestiones que inquietan a los administradores de colecciones digitales (Jones et. al. 2006):

¿Cómo pueden interactuar las interfaces de usuario con colecciones heterogéneas de objetos complejos de una manera simple e interoperable?

¿Cómo se puede asociar servicios y herramientas con objetos para proporcionar diferentes presentaciones o transformaciones de los objetos contenidos?

¿Cómo se puede asociar políticas de acceso a objetos específicos o a grupos de objetos?

¿Cómo se puede facilitar la preservación y el manejo a largo plazo de los objetos?

FEDORA no hace diferencia en el tipo de material que maneja. Además es flexible en cuanto permite a los implementadores del software diseñar sus propios modelos de contenido, para representar sus datos y sus requerimientos de presentación de la manera en que mejor les convenga o se adapte a sus necesidades esto por medio de los web services. Los objetos digitales y sus metadatos descriptores se ligan tratando a ambos elementos como un solo objeto compuesto; una facilidad que da FEDORA es que los metadatos de un objeto pueden estar almacenados localmente o en un servidor vía Web.

3.3. Propuesta de solución

Ambas soluciones de software se han considerado como opciones apropiadas para implementar RU para el prototipo de red de repositorios que se plantea en 3R, a fin de concretar las características de cada una en la tabla 1 se hace un resumen de las ventajas y desventajas de cada opción:

Tabla 1.- Comparativo entre DSpace y FEDORA

	Ventajas	Desventajas
DSpace	<p>Herramienta open source.</p> <p>Software listo para usarse.</p> <p>Es posible extender su funcionalidad con jsp's programados por quien lo implementa.</p> <p>Requerimientos mínimos de software adicional.</p> <p>Utiliza PostgreSQL como SMDB</p> <p>Lista para usarse, con posibilidad de configurarse para ajustarse a necesidades generales de repositorios de colecciones digitales: flujos de trabajo de aceptación y publicación de objetos, permisos de acceso a colecciones, restricción de acceso a nivel objeto, restricción de acceso a usuarios o grupos de usuarios, etc.</p> <p>Tiene programadas interfaces de administración, gestión y búsqueda y recuperación. Estas interfaces pueden ser modificadas por el instituto que lo implementa.</p> <p>Ampliamente utilizada alrededor del mundo.</p>	<p>Al estar lista para usarse, demanda de más trabajo dedicado a la adaptación para satisfacer necesidades particulares.</p> <p>La herramienta es ya un módulo, un núcleo, que necesita de un trabajo dedicado para ajustarse a necesidades muy particulares.</p>
FEDORA	<p>Herramienta open source.</p> <p>Cuenta con un trabajo exhaustivo de desarrollo orientado a tareas específicas (almacenamiento, gestión y preservación de acervos digitales) no dependientes del formato de los materiales.</p> <p>Es un núcleo de software con un modelo de funcionamiento abstracto que le permite ajustarse a requerimientos específicos.</p> <p>Cada vez cuenta con un mejor "soporte" que le permite resolver problemas en conjunto por medio de la comunidad fedora-users.</p> <p>Su funcionamiento se basa en los <i>web services</i>, que le permiten una interoperabilidad total con otros sistemas totalmente externos.</p> <p>Tiene puestos sus objetivos en la implementación de la web semántica.</p>	<p>Al ser un núcleo de software abstracto, no cuenta con interfaces programadas dedicadas a la administración, gestión y búsqueda y recuperación de contenidos.</p> <p>Por experiencia en el proyecto, y por lo leído en las listas de correo, todavía es común que surjan serios problemas en el funcionamiento general del sistema.</p> <p>Además de los que pueden salir en la instalación y la puesta en marcha. Hay detalles muy específicos (fallas en la ingesta de objetos, mal funcionamiento de interfaces programadas) que pueden obstaculizar el buen funcionamiento y entorpecer la labor del repositorio de quien lo implementa.</p>

Desde los inicios de este proyecto, se planteó que las dificultades más fuertes que se presentan en la implementación de repositorios no son de índole tecnológica, sino organizativa. Por lo anterior, consideramos que es mejor implementar algo que sea novedoso, tanto en la parte organizacional humana, como tecnológicamente hablando; es por esto que la principal apuesta está por FEDORA, aunque no se deja de lado el contemplar la utilización de DSpace como software para repositorios (o alguna otra herramienta compatible).

4. Modelo Conceptual de 3R

En los apartados anteriores se han abordado una serie de temas que tienen que ver con la definición y el desarrollo de RU, asimismo, se ha introducido al tema de las herramientas que se consideran adecuadas para implementar estos repositorios. Con estas ideas claras es momento de englobar el proyecto y delimitar un modelo conceptual para 3R.

Un modelo conceptual (GRIHO, 2007) proporciona información, de manera gráfica, acerca de qué es lo que hace un sistema y los mecanismos que implementa para llevar a cabo sus tareas y cumplir sus objetivos. Cabe recordar que el objetivo de 3R es la creación del prototipo de una red de repositorios de la UNAM, que permita mayor uso y visibilidad de la producción intelectual de los miembros de la comunidad. Por lo que en este apartado se describirá de qué forma se propone implementar el prototipo.

Inicialmente se presentará el modelo general, en el que se presentan de forma global todos los componentes del prototipo, posteriormente se detallarán los rubros de contenidos y de políticas, para finalmente detallar la arquitectura de un repositorio local, que corresponde a un RU dentro del modelo.

4.1. *Operación general*

3R es una red de RU, que ofrece servicios a dos niveles: uno local y otro global (figura 1). A nivel local (primer nivel), se encuentran los RU que son alimentados con la producción intelectual de las dependencias de la UNAM que se dedican a la docencia y a la investigación, y ofrecen servicios a una comunidad específica. A nivel global (segundo nivel) se encuentra una aplicación centralizada que ofrece una interfaz para la consulta de los RU locales, además de otros servicios generales que se explicarán más adelante.

El primer nivel tiene como función específica capturar, administrar y validar la información asociada a los objetos digitales depositados en cada uno de los repositorios de las Dependencias que integran el prototipo de red que plantea 3R, así mismo provee los metadatos para su cosecha por el servidor cache que reúne los metadatos de todos los repositorios asociados. A este nivel intervienen los siguientes elementos:

Actores:

- **Un administrador local que valida y organiza las colecciones del repositorio local.-** Tiene funciones de curador de las colecciones digitales, al tiempo que mantiene un orden, valida los metadatos asociados a los recursos y se encarga del control de las autorizaciones de acceso y depósito.
- **Usuarios locales.-** que como ya se describió en el apartado 2.2 pueden ser de dos tipos: **internos**, que son los encargados de alimentar el repositorio con objetos digitales producto de su labor de docente o de investigación y que a su vez pueden consultar el repositorio; **externos**, que consultan los contenidos del repositorio local.

Componentes:

- **Una interfaz de administración.-** Permite la gestión de los recursos digitales almacenados en el contenedor y la asignación de los permisos de acceso a diferentes niveles. También se encarga de vigilar el cumplimiento de las políticas internas del repositorio local
- **Servicios locales.-** Permiten las búsquedas y la recuperación de los objetos almacenados en el contenedor y permite la asociación del repositorio con otras comunidades de repositorios y con el portal central de 3R.
- **Repositorio local.-** Almacena al objeto digital en sí y sus metadatos asociados.

El segundo nivel centraliza el acceso y consulta pública de los metadatos cosechados en todos los repositorios que pertenecen a la Red y posibilita su intercambio con otros sistemas de información de otras organizaciones; este nivel tiene los siguientes elementos:

Actores

- **Usuario general.-** Persona que acude a 3R para hacer consultas a los contenidos de la Red, desde una interfaz centralizada o a través del protocolo OAI.
- **Administrador general.-** Puede establecerse como una estructura en forma de comité que se encargue de las distintas áreas de conocimiento y se encargue de la gestión y aplicación de las políticas generales que rigen al sistema. También tendrá a su cargo la supervisión técnica de todo el sistema en lo que se refiere al monitoreo del tráfico del acceso a la información y el cumplimiento periódico del respaldo y captura de la información proporcionada por los repositorios locales al caché.

Componentes

- **Portal de búsqueda y recuperación global.-** Provee una interfaz de acceso público que permita realizar búsquedas simultáneas en todos repositorios asociados del sistema, así como consultar información dentro de la estructura de comunidades y colecciones asociadas a estas. Es un proveedor de servicios.
- **Servidor caché fragmentado.-** Máquina encargada de cosechar periódicamente los metadatos asociados a los recursos almacenados en los repositorios universitarios y los pone a disposición de los usuarios a través del portal. Está fragmentado para permitir que usuarios y entidades externas puedan delimitar sus búsquedas en áreas de conocimiento específicas. Cada fragmento del caché corresponde a los metadatos específicos que provee cada repositorio local (se podría pensar que esta estructura del cache puede establecerse en función de criterios asociativos y temáticos de las familias de la información “asociación semántica de la información”).
- **Entidades externas.-** Entidades ajenas al sistema que pueden cosechar el conjunto de metadatos almacenados en los distintos fragmentos del servidor caché a través del protocolo OAI.
- **Servidor de preservación.-** Máquina que se encarga de cosechar tanto metadatos y objetos digitales de los repositorios. Este servidor no tendrá acceso público, su objetivo será garantizar la preservación a largo plazo de los objetos digitales (este componente sólo está sugerido para el nivel general del funcionamiento del sistema, el prototipo y la investigación no contemplan su implementación).

Funcionamiento

La interoperabilidad de ambos niveles esta dada por medio del protocolo para cosecha de metadatos OAI-PMH⁵, los repositorios locales del primer nivel se convierten en proveedores de datos que canalizan sólo metadatos referenciales hacia el cosechador OAI-PMH del segundo nivel, que los almacena y actualiza en el servidor caché, que tendrá funciones de proveedor de servicios a través del portal de acceso. Por medio de OAI-PMH se enlazará a los cosechadores externos al sistema.

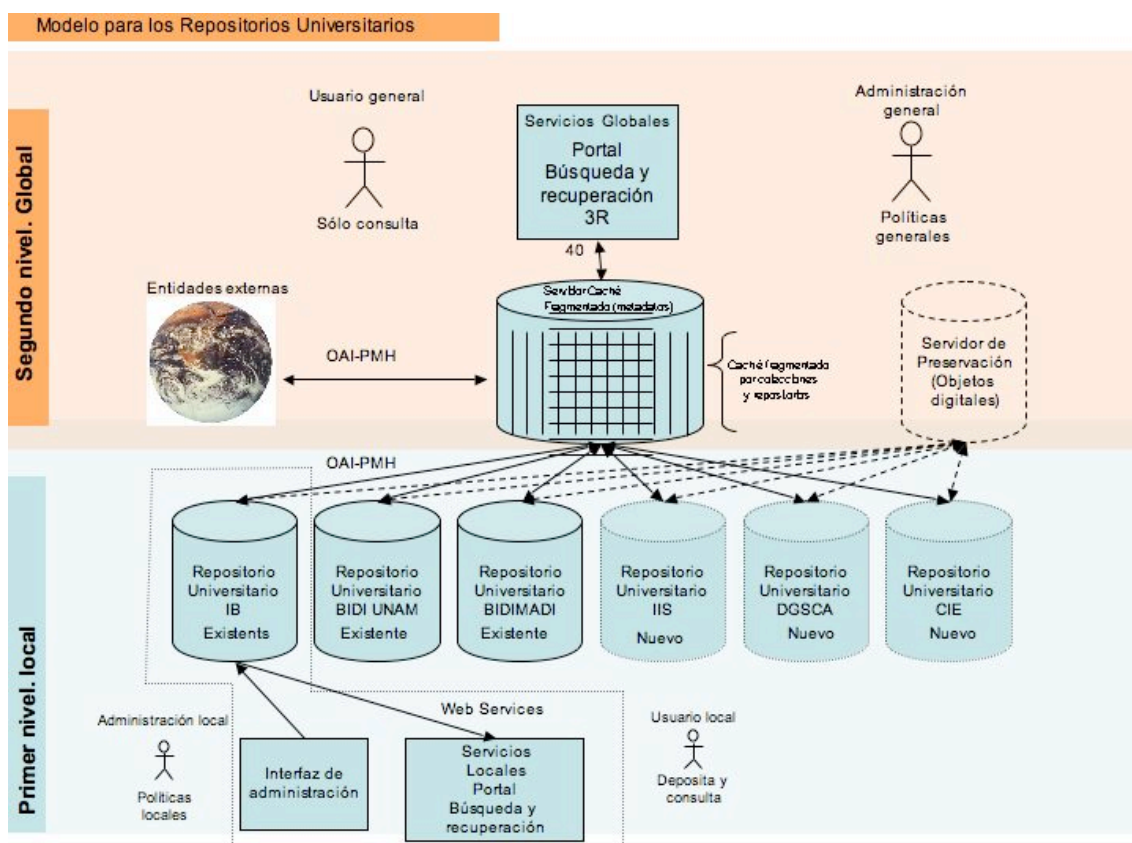


Figura 1. Modelo conceptual de 3R

4.2. Contenidos

Los contenidos son el valor primario de 3R, sin contenidos no tiene sentido la conformación de una red, por lo que se requiere definir el tipo de contenidos que se integrarán a los RU, un proceso de validación que certifique la coherencia de los metadatos y la correcta creación de colecciones, a fin de que los repositorios estén conformados adecuadamente y los servicios prestados con dichos contenidos sean estables y consistentes.

Dentro del contexto del modelo de 3R los contenidos se manejan como objetos digitales, ya que tendrán un tratamiento y una organización que los transforma en algo más que contenidos y se convierten en componentes de un repositorio que los administrará como objetos.

⁵ <http://www.openarchives.org/OAI/openarchivesprotocol.html>

4.2.1. Tipos de objetos digitales

Los tipos de objetos que se integran a los repositorios universitarios son una amplia gama de materiales y recursos académicos, distribuidos y producidos en sus distintas áreas; se presentan bajo distintas manifestaciones y múltiples formatos y expresiones. 3R trata de dar cabida a la mayor variedad posible de esta producción adecuando el sistema específicamente a la forma del contenido y las especificaciones de su formato original.

El manejo, formas de presentación y características técnicas de estos recursos estarán determinadas por las políticas adoptadas por las administraciones locales de los repositorios que integran 3R

La figura 2 muestra una lista de los tipos de objetos que pueden componer los RU. Esta lista trata de reflejar las distintas formas que toma la difusión de la investigación en la UNAM, la publicación editorial, la labor docente, y documentos que ayuden al funcionamiento administrativo de las dependencias que establezcan sus repositorios; los materiales de esta lista están contemplados independientemente de su formato, el depósito de los mismos será realizado de manera autónoma por los miembros autorizados de las dependencias a través de los mecanismos que proporciona el repositorio, quienes deberán proveer metadatos básicos (del esquema Dublin Core) asociados al objeto digital que depositan (figura 3). Los repositorios se organizarán de forma tal que deben contener tanto los metadatos como el objeto mismo.

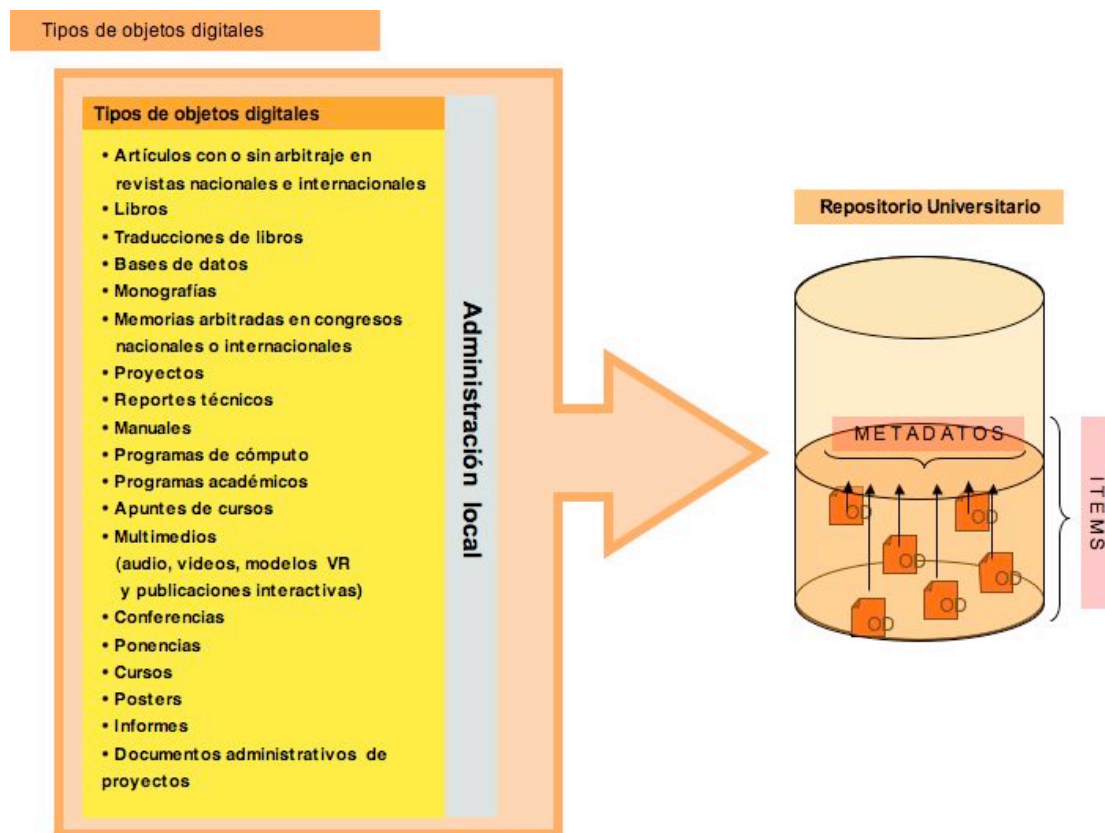


Figura 2. Tipos de objetos digitales.

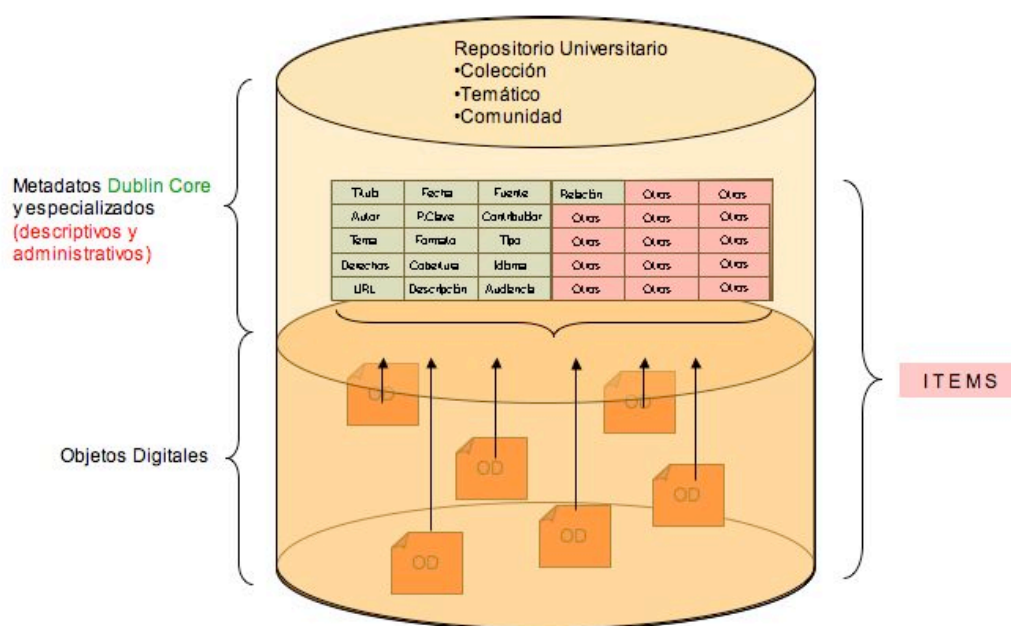


Figura 3. Tipos de objetos digitales (detalle)

4.2.2. Proceso de validación

Para asegurar la consistencia de la bases de datos y de los objetos depositados se debe contar con un proceso de validación (figura 4). En este proceso los metadatos de los objetos digitales depositados son examinados y validados por el administrador local para su depósito definitivo, se compone de los siguientes elementos:

- 1. Usuario local**, es el proveedor de objetos digitales y debe pertenecer a la comunidad académica del repositorio universitario, así como estar adecuadamente acreditado y autorizado por la administración del repositorio en el cuál deposita. Es importante recalcar que este actor del sistema es el responsable primario de los contenidos y los metadatos asociados de los recursos que deposita.
- 2. Administrador local**, es el encargado de mantener la estructura organizacional del repositorio (respecto a colecciones y comunidades), se encarga de la aplicación de las políticas de ingreso de materiales; encargado de la acreditación y autorización de otros usuarios y del proceso de validación de metadatos asociados a los objetos digitales. También se encarga de que los objetos sean persistentes e íntegros a lo largo del tiempo.
- 3. Repositorio Universitario** es la estructura de almacenamiento que contiene a los ítems aprobados para ponerlos a la disposición del público en general.

Para el ingreso de nuevos materiales a los repositorios y la validación de sus metadatos se proponen cualquiera de estos dos procedimientos:

a) Auto-depósito: en esta forma de ingreso, el autor de un recurso accede directamente al sistema mediante una interfaz de captura; ingresa los metadatos y el objeto digital. Antes de que el recurso pueda ser publicado y estar disponible, los metadatos, se someterán a un proceso de validación a cargo del administrador local del repositorio para asegurar su consistencia. Este proceso se compone de un contenedor temporal que retiene los objetos hasta su validación y que en caso de ser positiva, los objetos y sus metadatos son depositados de manera definitiva en el repositorio. En caso de que los metadatos del objeto no cumplan con alguno de los requisitos o políticas, el administrador local detendrá su incorporación y contactará con el usuario local depositante para que sean corregidos oportunamente.

b) Depósito a través de un administrador local: en esta forma de ingreso de materiales, el autor o propietario de los contenidos entrega el objeto digital y sus metadatos al administrador local para que éste se encargue de capturar los metadatos –al tiempo que los valida- e ingresar el objeto digital validado al repositorio.

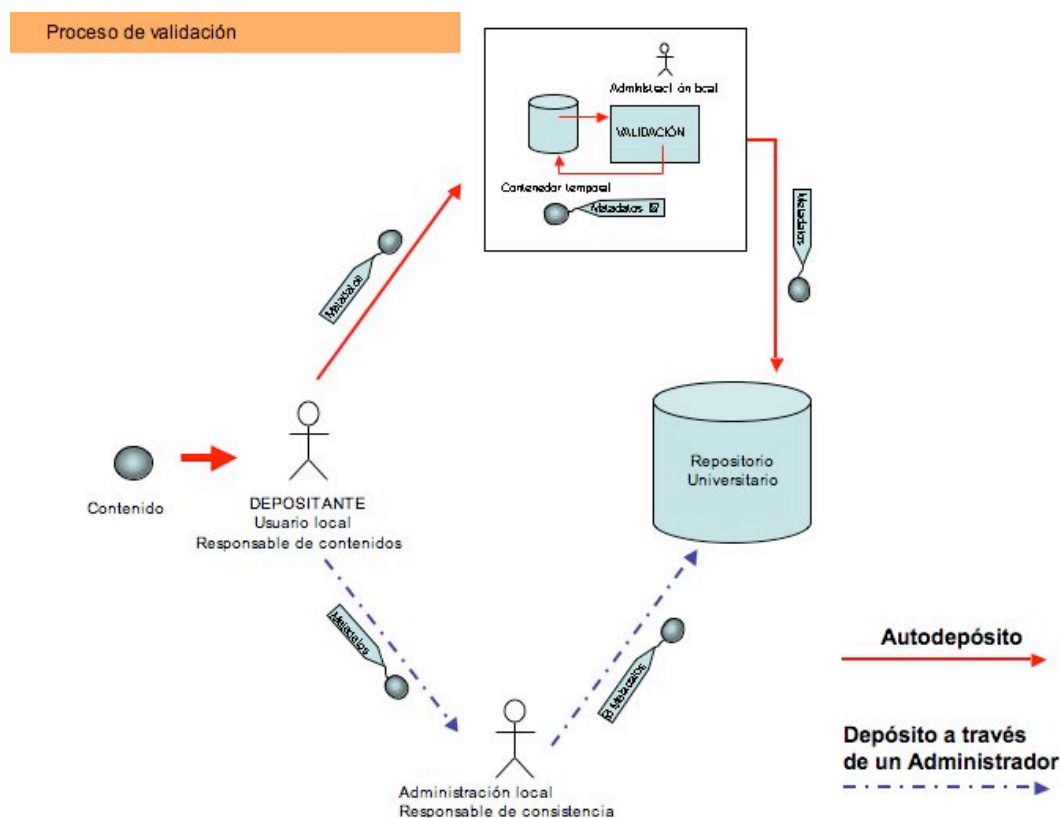


Figura 4. Proceso de validación de nuevos objetos.

4.2.3. Integración de colecciones

Al irse depositando contenidos en el repositorio se forman colecciones. Un mismo RU puede contener varias colecciones, y se hace de esta manera para agrupar contenidos que tienen alguna relación, usualmente temática. Las colecciones de objetos digitales se conforman de la siguiente manera (figura 5):

1. Inicialmente se encuentra el objeto digital tal cual será depositado y que puede presentarse en diferentes formatos. Dicho objeto es depositado por el usuario local y puede ser cualquiera de los tipos descritos en el apartado anterior.
2. Cuando el objetos de describe se transforma en un ítem, es decir, el objeto u objetos digitales ya dentro del sistema más sus metadatos asociados y validados por el administrador local.
3. Al irse acumulando los ítems se forman colecciones. Una colección es un conjunto de ítems agrupados bajo un contexto común y que el sistema asocia de manera interna para que sean consultados como un conjunto.
4. Para administrar una colección o un conjunto de ellas éstas se alojan bajo el concepto de un Repositorio Universitario. Un RU es la entidad de almacenamiento y administración que contiene a los ítems aprobados para ponerlos a la disposición del público en general. En su concepción más básica el RU esta constituido por objetos digitales y metadatos que funcionan como descriptores de dichos objetos. El estándar que se utiliza para describir objetos digitales es Dublin Core que tiene 16 elementos básicos y a los cuáles se les puede agregar un perfil de aplicación específico para el área de conocimiento que los implemente.

La gestión e integración de colecciones estarán a cargo del administrador local, a la colección y al repositorio. El envío inicial del objeto digital será responsabilidad del usuario local.

Integración de Colecciones

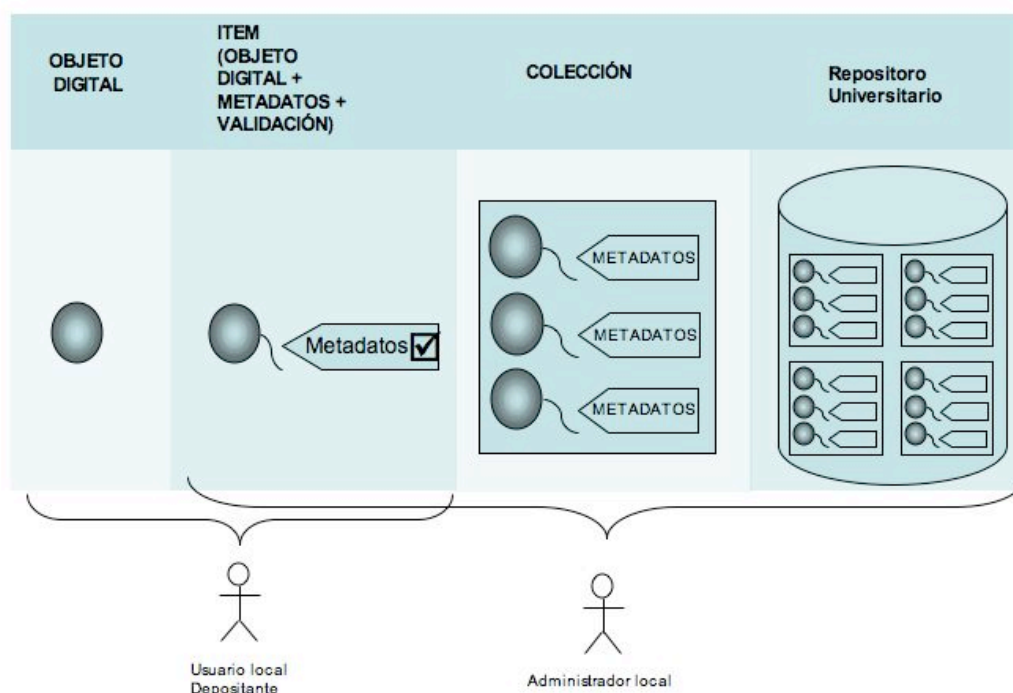


Figura 5. Integración de colecciones

4.3. Políticas

El proyecto 3R ha considerado establecer un marco mínimo para la formulación de políticas que regulen el funcionamiento los repositorios que lo integran, en esta etapa de desarrollo del proyecto estos principios deben ser tomados como un punto de partida y orientación inicial para formularse posteriormente y en consenso con los representantes de las comunidades académicas y responsables de los repositorios locales (Johnson, 2002). Es importante establecer los procedimientos y políticas para tener una expectativa clara sobre los roles y responsabilidades de las partes involucradas, en este sentido las políticas que 3R propone se avocan a tres aspectos generales: gestión de contenidos y colecciones, depósito y uso (figura 6).

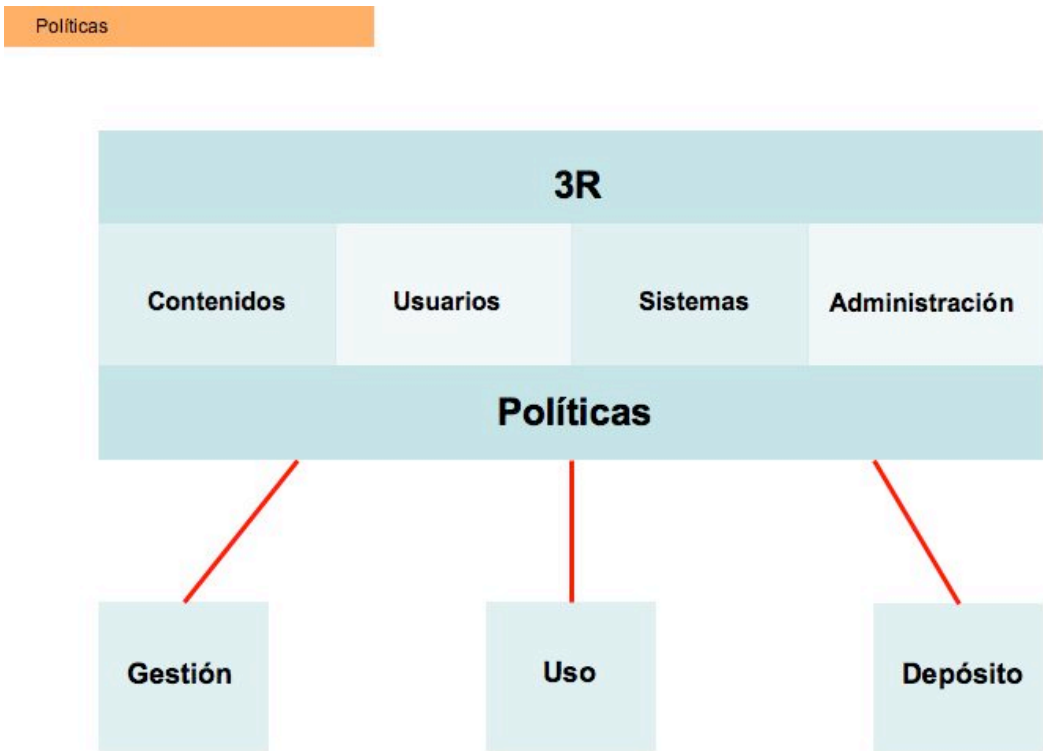


Figura 6. Políticas

4.3.1. Políticas sobre la gestión de contenidos y colecciones

Este aspecto es central para la constitución de cada RU ya que define el tipo de contenido válido, la forma en que está organizado, los lineamientos para su administración y los procedimientos para la validación de metadatos.

Cada dependencia definirá, sobre su RU:

- La estructura de sus colecciones.
- Los tipos y formatos de los documentos aceptados.
- Los procedimientos de revisión y aprobación para su depósito.

- Políticas de acceso a sus materiales.

Para los repositorios de 3R los contenidos son documentos que reflejen la investigación, docencia y difusión, producidos por los miembros académicos (profesores, investigadores y técnicos) pertenecientes a las Dependencias adscritas al prototipo.

Los documentos que integran las colecciones, deben cumplir con estas características:

- Ser producido (autor o coautor) por miembro(s) de la Dependencia.
- Resultar de actividades de investigación, difusión o enseñanza.
- Contar con los datos básicos para su identificación y clasificación.
- No ser efímero.
- Estar en formato digital.
- Estar completos.
- No tener problemas respecto a los derechos de propiedad intelectual y/o que estos no afecten a terceros.

Administración de colecciones:

- Cada dependencia definirá la estructura administrativa de sus colecciones.
- Contarán con un responsable o un grupo académico avalados por la Dependencia encargados de revisar y coordinar la aprobación de los documentos depositados en las colecciones.
- La dependencia es responsable de la calidad y validez de los contenidos de sus colecciones y tiene la facultad de aceptar, rechazar o retirar los depósitos de acuerdo a las requisitos exigidos para cada colección.
- La dependencia es responsable del otorgamiento de permisos para el depósito de documentos y asignación de metadatos.

Validación de metadatos

- Los repositorios 3R contendrán el recurso mismo y sus metadatos (ítem)
- La dependencia es responsable del proceso de validación (creación, edición y certificación) de los metadatos asociados a los documentos.
- Los repositorios 3R deben apearse al estándar mínimo de metadatos de Dublin Core simplificado, podrán así mismo, si así lo deciden, utilizar esquemas de metadatos adicionales.
- Deben utilizar un software recomendado por 3R que cumpla con el protocolo OAI-PMH para su interoperabilidad.

4.3.2. Políticas sobre procedimientos de depósito

La reglamentación de estos procedimientos facilita la gestión, integración y acceso a la información depositada en los repositorios universitarios. 3R debe proveer dos vías para la incorporación de documentos a sus repositorios: depósito directo (autodepósito) y el depósito mediado (Proberts, et. al. 2006).

- Podrán depositar publicaciones y documentos, los profesores, investigadores y técnicos académicos, avalados por las dependencias adscritas a 3R.
- Los responsables de los repositorios deberán proveer a los depositantes instrucciones claras y sencillas sobre el procedimiento, tipos y formatos del contenido a depositar. Así como proporcionar la opción de elegir la vía de entrega de los documentos: directamente (autodepósito) o mediante la entrega de los mismos a los servicios de documentación del repositorio para su proceso.
- El autor deberá ingresar los metadatos esenciales para que los documentos puedan iniciar su proceso de validación y aceptación.
- El autor deberá poseer los derechos de propiedad intelectual de su trabajo y conceder a la Dependencia, 3R y/o la UNAM el derecho no exclusivo de preservar y dar acceso a su documento a través de los servicios de búsqueda y consulta del repositorio, exceptuando los que no sean posibles por limitaciones de copyright u otras.
- El repositorio deberá proveer un servicio de ayuda y especificaciones sobre el marco legal de los derechos de propiedad intelectual (copyright) integrado al proceso de depósito.
- El repositorio deberá desarrollar servicios de valor agregado (listados, reportes de publicaciones, estadísticas de acceso y descarga, referencias, etc.) para los usuarios del repositorio, para fines informativos, de evaluación o administrativos (Rodrigues, 2006).

4.3.3. Políticas de uso

Estas políticas ayudan a la tener una idea concreta sobre el perfil de los usuarios y los tipos de contenidos que solicitan, determinan la forma en que se usan los documentos consultados y proporcionan indicadores de evaluación sobre el funcionamiento del repositorio.

- Es necesario el registro del usuario para tener acceso completo a la consulta del material.
- La administración del repositorio retiene el derecho de limitar el acceso al contenido, los metadatos o ambos. El acceso a los documentos puede ser libre,

restringido a usuarios y grupos autorizados o circunscrito únicamente a la comunidad universitaria.

- Es necesario que cada objeto cuente con una licencia de uso específica.
- El usuario final deberá aceptar los términos de un acuerdo o licencia general comprensiva para todos los contenidos depositados en el repositorio en el que se compromete a respetar lo establecido respecto al buen uso de los materiales publicados. El usuario puede descargar, salvar o imprimir los textos, resultados de búsqueda u otra información, sólo para uso personal, de investigación y docencia.

El éxito y estabilidad de los repositorios depende de manera decisiva de la operación de políticas que contribuyan gradualmente a promover la adhesión de nuevas comunidades, incrementar el número de usuarios (estimular el autodepósito) y aumentar el acervo de documentos. Paralelamente a las acciones de promoción y sensibilización sobre las ventajas de los repositorios de 3R es de esperarse que progresivamente las Dependencias, Centros de Investigación y la propia UNAM en su conjunto adopten formalmente una política institucional que considere obligatorio el depósito de su producción intelectual en los repositorios universitarios(Harnard, 2001; Pinfield, 2005).

4.4. *Arquitectura de un repositorio local*

La arquitectura de un RU o repositorio local se compone de cuatro elementos que interactúan en dos procesos distintos (figura 7):

- 1.- Repositorio universitario: Compuesto por una interfaz de servicios de administración y mantenimiento que permiten la funcionalidad del sistema y un contenedor de objetos digitales y sus metadatos. Un software de repositorios para la gestión de los mismos que permite la creación y uso de web services
- 2.- Web services: conjunto de herramientas con tecnología web que permiten la interacción entre distintos tipos de aplicaciones y servicios.
- 3.- Servicios locales: servicios que se ligan al repositorio para obtener información y procesarla de acuerdo a necesidades particulares de la comunidad de ese repositorio
- 4.- Portal de búsqueda y recuperación: permite el acceso a los metadatos cosechados de un conjunto de repositorios, partiendo desde una interfaz común y haciendo una sola petición de consulta por parte del usuario general.

El elemento central es el repositorio y sus servicios de administración y mantenimiento integrados a través del software específico para repositorios (FEDORA o DSpace).

Por un lado, el repositorio, a través de los web services, proporciona servicios adecuados a necesidades particulares de los usuarios de quien implementa el repositorio; por medio de estos servicios puede ligarse a distintos manejadores de bases de datos y/o portales temáticos que asocian parte de sus contenidos al depósito.

Por otra parte, este núcleo de servicios ofrece sus metadatos por medio del protocolo OAI-PMH al servidor caché o a cualquier sistema que lo requiera y de ahí brinda acceso a estos metadatos al usuario general por medio del portal de búsqueda y recuperación (servicios globales).

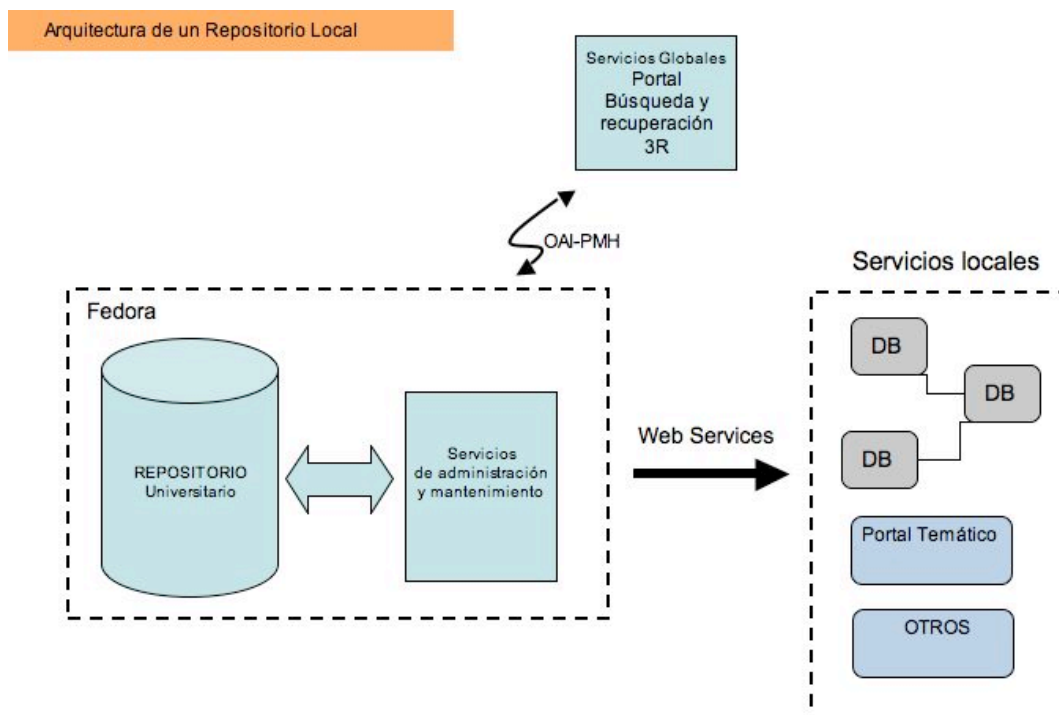


Figura 7. Arquitectura de un repositorio local

5. Conclusiones y trabajo futuro

Este segundo entregable del proyecto presento un modelo conceptual que sirve de punto de partida para la siguiente etapa que corresponde al desarrollo del prototipo. Durante esta parte de la investigación se detectó que las características de los repositorios de la UNAM no se adaptan adecuadamente, ni en organización ni en tecnología, a los tipos de repositorios que se encuentran comúnmente, ni en otras universidades ni en otros organismos, por lo que fue necesario hacer una definición propia para los repositorios universitarios, a partir de las definiciones de los tipos de repositorios comunes.

En el estudio de las herramientas de software se optó por seguir trabajando con dos opciones: FEDORA y DSpace, ya que hasta el momento de la implementación y de las pruebas se podrá determinar cuál de ellos es realmente más potente y cubre los requisitos para los RU que se plantean en la UNAM. Hasta ahora, hay una tendencia por el uso de FEDORA, por ser un software poderoso y robusto, sin embargo, debido a que su uso aún no es tan extendido, se requiere de más pruebas para tomar una decisión al respecto.

Con el análisis del modelo y la arquitectura tecnológica planteada se confirma que el principal reto no es la parte tecnológica, ya que hay soluciones confiables que se pueden utilizar y hay protocolos que ya realizan las funciones que se requieren para la interconexión y el procesamiento de metadatos; la parte de organización, la logística de operación de los repositorios y de la red, así como el asumir las políticas en la práctica, será la parte que conlleve más esfuerzo.

Las siguientes actividades del proyecto estarán encaminadas a la tercera fase, que consiste en el desarrollo de las aplicaciones para implementar la red de repositorios. Para ello, es necesario, primero, permear y discutir el modelo conceptual entre los repositorios que se seleccionaron en la etapa I, a fin de recibir retroalimentación y evaluar posibles ajustes. Segundo, se dará inicio a las pruebas con las herramientas de software y se realizarán las modificaciones pertinentes. También será labor importante el que cada repositorio del prototipo comience a trabajar en sus políticas locales y se tengan talleres para debatirlas y obtener políticas un tanto homogéneas, aunque no similares ya que cada repositorio local tendrá sus propios requerimientos.

6. Referencias

GRIHO. "Modelo mental y modelo conceptual".

<http://griho.udl.es/mpiua/modelosmyc.htm>

Harnard, S. (2001). "The Self-Archiving Initiative." *Nature* 410. 1024-1025.

<http://www.ecs.soton.ac.uk/~harnad/Tp/nature4.htm>

Johnson, Richard K. (2002). "Institutional repositories: partnering with faculty to enhance scholarly communication"; *D-Lib Magazine*, 2002, 8 (11)

<http://www.dlib.org/dlib/november02/johnson/11johnson.html>

Jones, R., Theo A. and MacColl, J. (2006), *The Institutional Repository*, Oxford, Chandos Publishing, 250 pp.

Pinfield, Stephen (2005), "A mandate to self archive? The role of open access institutional repositories" *Serials* 18(1), March, 2005

http://eprints.nottingham.ac.uk/archive/00000152/01/mandate_to_archive.pdf

Proberts, Steve y Jenkins, Celia (2006). "Documentation for Institutional Repositories", *Learned Publishing*, 19 (1), 57-71

<http://magpie.lboro.ac.uk/dspace/bitstream/2134/782/1/lppaper+-+final.pdf>

Rodrigues, Eloy (2006). "Poniendo en práctica el acceso libre a la literatura científica : el repositorio institucional y la política de auto-archivo de la Universidade do Minho", *Jornadas das Bibliotecas del Consejo Superior de Investigaciones Científicas (Csic)*, 3, Madrid, España

<http://repositorium.sdum.uminho.pt/handle/1822/5252>

7. Bibliografía

Callan, P. (2004). The development and implementation of a university-wide self-archiving policy at Queensland University of Technology (QUT): insights from the frontline.

http://www.arl.org/sparc/meetings/ir04/presentations/callan_files/callan.pdf

GRIHO. "Modelo mental y modelo conceptual".

<http://griho.udl.es/mpiua/modelosmyc.htm>

Harnard, S. (2001). "The Self-Archiving Initiative." *Nature* 410. 1024-1025.

<http://www.ecs.soton.ac.uk/~harnad/Tp/nature4.htm>

Johnson, Richard K. (2002). "Institutional repositories: partnering with faculty to enhance scholarly communication"; *D-Lib Magazine*, 2002, 8 (11)

<http://www.dlib.org/dlib/november02/johnson/11johnson.html>

Joint, Nicholas (2006). "Institutional Repositories, Self-Archiving and the Role of the Library." *Library Review* 55(2), 81-84.

<http://www.emeraldinsight.com/Insight/viewContentItem.do?contentType=Article&hdAction=lnkhtml&contentId=1545768>

Jones, R., Theo A. and MacColl, J. (2006), *The Institutional Repository*, Oxford, Chandos Publishing 250 pp.

Lagoze, C., Payette, S. (1998) "Flexible and Extensible Digital Object and Repository Architecture".

<http://www.cs.cornell.edu/payette/papers/ECDL98/FEDORA.htm>

Lynch, Clifford A.(2003). "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age." *ARL Bimonthly Report [Revista electrónica]* 226.

<http://www.arl.org/newsltr/226/ir.html>

Payette, S., Staples, T. (2002) "The Mellon Fedora Project: Digital Library Architecture Meets XML and Web Services".

<http://www.fedora.info/documents/ecdl2002final.pdf>

Pinfield, Stephen (2005), "A mandate to self archive? The role of open access institutional repositories" *Serials* 18(1), March, 2005

http://eprints.nottingham.ac.uk/archive/00000152/01/mandate_to_archive.pdf

Proberts, Steve y Jenkins, Celia (2006). "Documentation for Institutional Repositories", *Learned Publishing*, 19 (1), 57-71

<http://magpie.lboro.ac.uk/dspace/bitstream/2134/782/1/lppaper+-+final.pdf>

Publisher Copyright Policies and Self Archiving

<http://www.sherpa.ac.uk/romeo.php>

Rodrigues, Eloy (2006). “Poniendo en práctica el acceso libre a la literatura científica : el repositorio institucional y la política de auto-archivo de la Universidade do Minho”, *Jornadas das Bibliotecas del Consejo Superior de Investigaciones Científicas (Csic)*, 3, Madrid, España
<http://repositorium.sdum.uminho.pt/handle/1822/5252>

8. Anexos

Estas tablas son una evaluación de las principales características de los tres modelos de repositorios universitarios planteados por 3R.